**DOCUMENT RESUME**

ED 157 497 IR 005 756

AUTHOR Carroll, John M.; Tague, Jean M.
TITLE Community Information Centers and the Computer.
PUB DATE 75
NOTE 9p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Community Information Services; *Comparative Analysis; Computers; Data Bases; Human Services; *On Line Systems; Performance; *Relevance (Information Retrieval)
IDENTIFIERS Ontario

ABSTRACT

Two computer data bases have been developed by the Computer Science Department at the University of Western Ontario for "Information London," the local community information center. One system, called LONDON, permits Boolean searches of a file of 5,000 records describing human service agencies in the London area. The second system, called INDEX, consists of the same information classified into 45 subject-oriented files. Users retrieve a file by subject keyword and browse through it. Online retrieval using the two computer-based files and manual searching using the telephone directory were compared in an experiment carried out by students at the UWO School of Library and Information Science. Systems were evaluated on the basis of answers to a set of 16 questions, both with respect to the time taken to obtain a correct answer and the number of satisfactory or correct answers produced. Four analyses of variance were applied to the data obtained: effect of question order on proportion of correct answers, comparison of the three systems in terms of time to reach satisfactory answers, time to reach an answer when location of the information in the data base was known, and proportion of questions correctly answered. Results indicate that the greatest potential for the computer lies in the production of agency lists and updated manual indexes rather than in direct online retrieval. (Author/JPF)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED157497

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

# COMMUNITY INFORMATION CENTERS AND THE COMPUTER

by

John M. Carroll

Jean M. Tague

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John M. Carroll

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."



IR005756

# COMMUNITY INFORMATION CENTERS AND THE COMPUTER
# (L'ORDINATEUR ET LES RENSEIGNEMENTS POPULAIRE)

John M. Carroll and Jean M. Tague
University of Western Ontario
London, Ontario, N6A 3K7

## ABSTRACT

Two computer data bases have been developed by the Computer Science Department at the University of Western Ontario for the local community information center, Information London. One system, called LONDON, permits Boolean searches of a file of 5,000 records describing human service agencies in the London area. The second system, called INDEX, consists of the same information classified into 45 subject-oriented files. Users retrieve a file by subject keyword and browse through it. On-line retrieval using the two computer-based files and manual searching using the telephone directory were compared in an experiment carried out by students at the UWO School of Library and Information Science. Results indicate that the greatest potential for the computer lies in the production of agency lists and updated manual indexes rather than in direct retrieval.
(Le Department de l'Informatique de l'Universitie de Western Ontario a fait deux banques des donnés pour Renseignements London, un centre locale des renseignements populaire. La premiére est une fiche de 5,000 dossiers, chaqu'une de laquelle décrit quelque service humanitaire. On entre dans cette fiche avec les combinaisons logiques des codes de classification.
La deuxième consist de 45 de fiche qui ont la meme information classifiée par sujet. On entre dans cette fiche avec le titre du sujet seulement. Nous avons fait une comparison experimentale entre les deux systemes et aussi avec l'usage de la directoire telephonique. Dans cette experimentation les étudients de l'Ecole des Sciences Bibliothêques ont cherchés les fiches. La resultat indique que probablement le futur le plus grand pour l'ordinateur est à produire les directoires à main speciales; ce' n'est pas peut-être a chercher à la ligne.)

## COMMUNITY INFORMATION

Since 1972 the computer science department has been cooperating with Information London to determine in which of their areas of concern computer assistance would be effective and how such assistance can best be rendered. Information London is a voluntary organization that provides a free question answering service, principally by telephone, to persons in the London area.

One area we explored was the question-answering process. The mode of computer assistance initially investigated was that of storing information within a time-sharing computer system in such a way that it could be recalled at a terminal in the offices of Information London in response to client inquiries.

We implemented two retrieval systems which differed substantially both in design philosophy, file maintenance procedures, and retrieval strategy. This paper describes and discusses an operational comparison under controlled conditions among these two systems and manual information retrieval based upon a printed directory.

### THE SYSTEMS

The first system (LONDON) consisted of a single master file containing 5,000 records. The records were each 300 characters long and each gave the name, address, telephone number, and a locally designed classification code pertaining to some agency or business, or institution offering some human service.

The data bank was constructed on an *a priori* basis by culling from published local directories information that was presumed to answer anticipated client inquiries. Our perception of the information needs of clients was obtained from a study of past client-contact reports. One of these reports is made after each call received by Information London.

Access to the LONDON system is gained by prescribing set intersection, union, or negation of descriptors or digits of the classification number. The latter specify the human services provided by each subject organization. The former consist of keywords derived from the name field of the master file record.

Internally the master file records are ordered on their classification numbers which have had computer-generated tie-breaking digits appended to them. An inverted file is generated for each description or classification digit. The system uses 3,000 descriptors and 40 classification digits.

System INDEX represents a totally different approach. It is based upon a collection of 45 files each with an average of 100 records. Record format is completely unstructured and frequently contains narrative information as well as name, address, and telephone number of a

human-service organization.

The data bank was constructed on an a posteriori basis. The information in it was developed for the most part by the information officers of Information London in response to client inquiries and recorded manually in note format at the time of client contact. From time to time these handwritten notes are collated by the agency secretary, typed, reproduced and distributed among all the information officers. Each file has a title descriptive of a generic attribute shared by the organizations listed in it such as ethnic groups, arts and crafts, child care etc. Preparation of the data bank simply involves typing the information at a computer terminal rather than a typewriter.

In system INDEX we aggregate these files into groups of nine and assign group names that embrace all the files such as: welfare and benefits, recreation, consumer services etc.

When a user addresses the system, he is presented with a sequence of displays each consisting of the file names belonging to a group. The user selects the file he wants to see by depressing the key corresponding to a number belonging to that file. (In this way we are simulating use of a touch-sensitive display panel.)

The system then calls the selected file and displays a list of keywords. These keywords consist simply of words selected from among the first significant words of the lines in the file. Having thus made entry to the file, the user is able to browse through it in case the record retrieved fails to satisfy his needs.

File Maintenance

The experiment described in this paper will focus on comparing retrieval operations with LONDON and INDEX. None the less important, however, are file maintenance operations.

The LONDON file with its fixed-format records requires a formal updating procedure with the usual use of coding forms, keypunching and verification for data entry and a batch run on the computer to create a new master file and all the necessary inverted files. This has proved to be so expensive and inconvenient that there is a strong tendency to forego making minor changes to records over a period of time with the result that the information in the files becomes outdated and unreliable.

The updating of the INDEX file requires no special skills beyond typing and the use of a subset of text-editor commands. Best of all, the updating procedure is already embedded into the operating routine of the information centre and the only change is that files that

have to be created anyway are now in the first instance created in machine-sensible format rather than in hard copy. Derivative benefits accrue in the form of making it possible to insert changes without retyping a page and the ability to make copies of files or portions of files for the convenience of clients and staff.

The fixed-format of system LONDON is seen as a problem by some information officers who find that additional information regarding some people-serving organizations is necessary but may be of a kind one would not wish to become a standing element of every record.

System INDEX, for example, permits us to note that the St. Vincent de Paul society provides a free soup-sandwich-and-coffee lunch for the destitute; or that Mission Services has an apartment available for a mother and children who have to flee on short notice from a potentially dangerous family situation. Not only can such miscellaneous information be accommodated by INDEX, it can also be added whenever it becomes known without requiring an expensive system update and cocomittant interruption of routine operations.

On the positive side, LONDON permits through use of boolean operators the preparation of lists of organizations sharing attributes whose coordination was unforeseen at the time the file was constructed. Although some users profess to dislike the use of numerical classification codes when first encountered, we find, in fact, that most eventually come to use them and with some they become the entry points of choice.

INDEX, on the other hand, appears to suffer from its inability to manipulate records on a unit basis and the paucity of file access points.

## THE EVALUATION

During the fall of 1974, the authors and three students from the University of Western Ontario School of Library and Information Science made several visits to Information London, to determine the effectiveness of the LONDON and INDEX computerized retrieval systems. It was found that little actual use was taking place, partly as a result of physical inconvenience--the terminal was situated in a different room from the telephone answering desks--but also as a result of the disinclination of the Information London personnel to use the computerized system. It was therefore decided to carry out an evaluation of the two systems, to pinpoint the problems they presented to the nonspecialist user and to determine improvements which could be made.

In the following winter, 1975, term, a class of 45 SLIS students compared the effectiveness of retrieval from three bases: LONDON, INDEX, and, as representative of the manual approach, the London-

St. Thomas Telephone Directory. Systems were evaluated on the basis of answers to a set of 16 questions, both with respect to the time taken to obtain a correct answer and the number of satisfactory or correct answers produced.

Design of Experiment

The 16 questions were selected from the back files of Information London. Since the primary purpose of the Information London data base is to provide factual information (name, address, telephone number, contact person) on human service agencies, the questions all required answers in this form rather than value judgements or counselling. Each question could in fact be answered by any of the three systems, thus the test was of usability rather than the covereage of the data bases.

Questions were assigned randomly to students and to systems in such a manner that each student searched three questions on each of the three systems, with no student doing a particular question on more than one system. The original intention had been to repeat each question seven times on each system. However, for various reasons some students were unable to complete the assignment, so that the actual number of repetitions varied from five to seven.

Students recorded the time spent in answering each question and whether or not an agency at least potentially capable of providing the information was located. Prior to the test period they were given a description of the two computerized systems and permitted, for a period of approximately one week, to familiarize themselves with the terminal access. Thus, they might be classed as informed nonspecialist users. As a control, the authors determined the time needed to obtain answers on the two computerized systems when the location of the answer was known, as it would be to an experienced specialist user.

Analysis of Data

Four analyses of variance were applied to the data:

(1) Comparison of the three systems in terms of time in minutes to reach satisfactory answers.

(2) Comparison of the three systems in terms of time to reach an answer when location of the information in the data base was known.

(3) Comparison for the three systems of proportion of questions correctly answered.

(4) Effect of question order (first, second or third question answered on a system) on proportion of correct answers.

The mean times to locate an answer for the three systems are as follows:

| System | Time(min.) |
|---|---|
| LONDON | 7.80 |
| INDEX | 13.02 |

| System | Time(min.) |
|---|---|
| Manual | 2.97 |

The ANOVA and a subsequent Newman-Keuls analysis showed significant differences among all pairs of means at the .01 level.

Some comments should be made on the outcome. The apparent overwhelming superiority of the manual system is to some extent a result of the greater familiarity of the students with that system. Also, although no comparison was made of the degree of relevance of the answers, the lists of agencies produced by the computerized systems were, in general, more complete. Finally, it is suspected that the students underestimated the time required to record the answer in the manual search.

Much of the time spent at the terminal was learning time. This problem is apparent from the results of the second analysis in which average time to obtain an answer when the location of the answer is known was compared for the three systems. The means were as follows:

| System | Time(min.) |
|---|---|
| LONDON | 2.88 |
| INDEX | 3.94 |
| Manual | 2.97 |

The second analysis showed no significant differences among the three systems at the .05 level.

The over-all numbers and proportions of questions answered satisfactorily were as follows:

| System | No. of Trials | No. of Answers | Proportion |
|---|---|---|---|
| LONDON | 107 | 73 | .68 |
| INDEX | 108 | 75 | .69 |
| Manual | 106 | 86 | .81 |

Again, greater familiarity with the phone book led to its apparent superiority in locating answers. However, differences among the three systems were not significant at .05 level.

The final analysis, which tested for a learning effect in repeated use of the same system, produced nonsignificant results.

Conclusions

In general, students preferred the Boolean query capability of LONDON to the browsing, file-display capability of INDEX. INDEX provides no mechanism to move from file to file in any desired sequence. However, students liked the "Professional" mode of INDEX after the first two or three trials, and suggested a similar mode be implemented on LONDON. The extensive dialog and prompting facilities on LONDON were found to be cumbersome.

For the nonspecialist user, for the man-in-the-street, computer-

based retrieval systems do not seem to be practical. The Information London user must understand thoroughly the structure of the classification system, the method for formulating queries, and the basis for inclusion in the keyword index in order to use LONDON effectively, and few such users will have either the time or the inclination for such a study. INDEX appeared by its very nature to be an inefficient retrieval tool, although of some value in browsing. Generally speaking, neither the users nor the personnel of a community information center such as Information London are computer-oriented, as might be the case with users of CAN-OLE, for instance. Community information center personnel feel more at ease with books than terminals and, because of the personal counselling aspect of their service, prefer to obtain answers by direct contact with agencies whenever there are problems or ambiguities. Only purely factual questions can be handled by a computerized data base, both because of the nature of the information and for reasons of confidentiality, and the tests seem to indicate that factual information is most rapidly obtained from an up-to-date manual index.

One should not conclude, however, that the computer has no function in a community information center. In the test, both systems, LONDON and INDEX, proved useful in recording or listing lengthy answers. The optimal potential for computer use would seem to lie in the periodic production of up-dated manual indexes, organized in a variety of formats to provide multiple access points to the data base. In this way, the rapid updating facility of the computer and the browsing facility of the printed book are merged to provide a retrieval mechanism acceptable to both users and administrators.

Any such utilization will, of course, be acceptable only if the computerized data base is thoroughly debugged, maintained, and updated on a regular basis by professional computer personnel. Such experts are not normally found in community information centers. Thus, any proposal for computerization of a community information center should take into account the need for a continuing source of funding to maintain the system.